

Criminal Justice and Behavior

<http://cjb.sagepub.com/>

Violent Recidivism of Mentally Disordered Offenders : The Development of a Statistical Prediction Instrument

GRANT T. HARRIS, MARNIE E. RICE and VERNON L. QUINSEY

Criminal Justice and Behavior 1993 20: 315

DOI: 10.1177/0093854893020004001

The online version of this article can be found at:

<http://cjb.sagepub.com/content/20/4/315>

Published by:



<http://www.sagepublications.com>

On behalf of:



International Association for Correctional and Forensic Psychology

Additional services and information for *Criminal Justice and Behavior* can be found at:

Email Alerts: <http://cjb.sagepub.com/cgi/alerts>

Subscriptions: <http://cjb.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Downloaded from cjb.sagepub.com at JUSTICE DEPARTMENT on June 29, 2010

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

VIOLENT RECIDIVISM OF MENTALLY DISORDERED OFFENDERS

The Development of a Statistical Prediction Instrument

GRANT T. HARRIS

MARNIE E. RICE

Mental Health Centre, Penetanguishene

VERNON L. QUINSEY

Queen's University, Kingston

Multivariate techniques were used to derive and validate an actuarial instrument for the prediction of violent postrelease offenses by mentally disordered offenders. The 618 subjects were a heterogeneous group of men who had been charged with serious offenses. Approximately half had been treated in a maximum security psychiatric institution and the rest had been briefly assessed prior to imprisonment. The actuarial instrument consisted of 12 variables and significantly predicted violent outcome in each of five subgroups. The instrument's practical application and its use in clinical appraisals of dangerousness are discussed.

The law requires that predictions of dangerousness be made in many circumstances. Courts, parole boards, psychiatric review boards, and individual clinicians are called on to decide, at least partly

AUTHORS' NOTE: *The authors thank C. Cormier, C. Lang, V. Bell, L. Koeshkerian, C. MacKnight, and S. Robinson for assistance in data collection and N. Zoe Hilton for helpful comments on an earlier version. This research was funded by grants from the Ontario Mental Health Foundation and the Ontario Ministry of Health and a contract from the Mental Health Facilities Branch of the Ontario Ministry of Health. Correspondence should be addressed to Grant T. Harris, Research Department, Mental Health Centre, Penetanguishene, Ontario, Canada, LOK 1P0.*

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 20 No. 4, December 1993 315-335
© 1993 American Association for Correctional Psychology

for the protection of the public, who is to be incarcerated or institutionalized and for how long. As is the case in virtually every decision-making situation that has been examined, statistical prediction has been found to be more accurate than expert clinical judgment (Gottfredson, 1987). Over the past several decades, there have been many attempts to derive statistical or actuarial methods to predict the level of risk that offenders present to the community. Most of the instruments have been designed to predict criminal recidivism of any kind. It was not until the 1970s and 1980s, however, that statistical prediction instruments were systematically applied in practice and, with the use of these instruments, the predictive accuracy of general criminal recidivism is now routinely in the 60% to 80% range (Andrews, 1989).

Such instruments have been far less successful, however, at predicting violent recidivism (Holland, Holt, Levi, & Beckett, 1983; Nuffield, 1982). There is also general agreement that clinical prediction in this area is extremely poor (Steadman, 1987). At the same time, it is violence that the public and decision makers most want to predict and prevent. Therefore, any progress toward the development of a useful instrument for predicting violent recidivism would be most welcome, and recent reviews (e.g., Monahan & Steadman, *in press*) have suggested several methodological improvements for future research.

One of the biggest problems in the area of prediction of violence has been the problem of low base rates (Quinsey, 1980; Steadman, 1987). Typically, the likelihood of violent recidivism in a sample of offenders is so low that it is hard to improve on the prediction that no one will commit a violent offense (Steadman, 1983). Thus, for example, in a study of over 1,000 Canadian federal inmates, the predictive abilities of three statistical devices for identifying violent recidivists were so poor that none of the instruments could be recommended (Nuffield, 1982).

In a series of follow-up studies recently completed in Ontario, however, rates of violent recidivism ranging from 16% to 77% have been reported (Harris, Rice, & Cormier, 1991; Quinsey & Maguire, 1986; Rice & Harris, 1992; Rice, Harris, & Cormier, 1992; Rice, Harris, Lang, & Bell, 1990; Rice, Harris, & Quinsey, 1990; Rice, Quinsey, & Harris, 1991; Rice, Quinsey, & Houghton, 1990). In some of these studies (Rice et al., 1992; Rice, Harris, Lang, & Bell, 1990; Rice, Harris, & Quinsey, 1990; Rice et al., 1991), the usefulness of a

large number of variables suggested by the literature on the prediction of dangerousness (Monahan, 1981; Quinsey, 1984) has been examined in an attempt to identify which variables, taken separately and in combination with other variables, best predict violent recidivism among mentally disordered offenders and their nonmentally disordered criminal counterparts.

The purpose of the present study was to combine some of these samples and their common variables to obtain a large data set for the construction of a statistical instrument for the prediction of violent criminal recidivism. The value of such an instrument, if sufficiently accurate, would lie in its application to new cases. Individuals receiving low scores on such an instrument would be candidates for relatively short sentences, early release, or less intensive community supervision. On the other hand, offenders with high scores would be candidates for longer sentences, preventive detention, or intensive community supervision.

METHOD

SUBJECTS

The subjects have been described in detail elsewhere (Rice et al., 1992; Rice, Harris, Lang, & Bell, 1990). Briefly, the initial subject pool was comprised of two groups: (a) 371 men (of whom 332 eventually had an opportunity to recidivate; see below) admitted for treatment to a maximum security psychiatric institution in Ontario between 1965 and 1980 and (b) 324 men (of whom 286 had an opportunity to recidivate) who had been admitted only for a brief pretrial psychiatric assessment. Each of the men in the second (assessed) group was matched to a man in the first (treated) group according to the following criteria: (a) each had received the same criminal charge for his index offense, (b) both had the same score (within 20%) on a measure of their frequency and severity of both past violent and nonviolent criminal activity (Akman & Normandeau, 1967), (c) both were the same age (within 1 year), and (d) their index offenses occurred within 12 months of each other. Background and offense data for the subjects are shown in Table 1. These two groups

TABLE 1: Comparison Between Subjects Who Recidivated Violently and Those Who Did Not

| <i>Variable</i> | <i>Subjects</i> | | <i>t or χ^2</i> |
|--|---|--|---------------------------------|
| | <i>Recidivists^a</i> (<i>n</i> = 191) | <i>Nonrecidivists^a</i> (<i>n</i> = 427) | |
| Childhood history | | | |
| Highest school grade | 8.0 (2.5) | 9.0 (2.4) | 4.67* |
| Elementary school maladjustment | 2.6 (1.2) | 1.8 (1.1) | 7.88* |
| Teen alcohol abuse score | 1.5 (1.1) | 1.2 (.93) | 2.55 |
| Socioeconomic status | 318 (139) | 310 (161) | n.s. |
| Childhood aggression | 3.2 (1.9) | 2.3 (1.7) | 6.06* |
| Behavior problems | 3.5 (2.6) | 2.0 (5.3) | 3.71* |
| Suspended or expelled (%) | 25 | 14 | 11.68* |
| Arrested under age 16 (%) | 42 | 19 | 38.21* |
| Separation from parents under age 16 (%) | 55 | 28 | 38.17* |
| Parental crime (%) | 13 | 6 | 6.82 |
| Parental psychiatric history (%) | 16 | 12 | n.s. |
| Parental alcoholism (%) | 51 | 37 | 9.29 |
| Adult adjustment | | | |
| Longest employment (mo) | 30 (83) | 66 (127) | 3.44* |
| Admissions to corrections | 1.7 (2.2) | .81 (1.8) | 5.31* |
| Psychiatric admissions | 1.1 (2.0) | 1.3 (2.4) | n.s. |
| Alcohol abuse score | 2.2 (1.6) | 1.8 (1.6) | 3.16 |
| Impulsivity score | 3.1 (1.8) | 2.3 (1.7) | 4.90* |
| Property offense history | 7.2 (14.9) | 2.9 (7.8) | 4.62* |
| Violent offense history | 3.3 (7.7) | 1.9 (6.4) | 2.41 |
| Never married (%) | 70 | 50 | 20.23* |
| Previous violent offense (%) | 35 | 23 | 10.94* |
| Ever fired (%) | 34 | 34 | n.s. |
| Escaped from an institution (%) | 24 | 11 | 17.69* |
| Failure on prior conditional release (%) | 43 | 20 | 36.69* |
| Lived alone (%) | 46 | 37 | 4.10 |
| Index offense | | | |
| Age at index offense | 23.5 (7.2) | 29.7 (11.9) | 6.63* |
| Victim injury | 3.6 (2.4) | 4.4 (2.2) | 4.04* |
| Seriousness of index offense | 19.4 (19.3) | 22.1 (20.9) | n.s. |
| Violent offense (%) | 76 | 85 | 7.46* |
| Victim knew offender (%) | 28 | 42 | 11.11* |
| Female victim (%) | 39 | 51 | 7.07* |
| Weapon used (%) | 61 | 70 | 4.89 |
| Sexual motive (%) | 14 | 7 | 8.24 |
| Alcohol involved (%) | 51 | 43 | n.s. |

continued

TABLE 1: continued

| Variable | Subjects | | |
|---|---------------------------------------|--|---------------|
| | Recidivists ^a (n = 191) | Nonrecidivists ^a (n = 427) | t or χ^2 |
| Assessment results | | | |
| IQ | 97 (17) | 99 (14) | n.s. |
| Level of Supervision Inventory | 19.8 (8.1) | 15.5 (8.5) | 5.81* |
| Psychopathy Checklist | 21 (8.7) | 14 (9.0) | 8.47* |
| Elevation on MMPI Scale 4 (%) | 30 | 22 | 5.59 |
| DSM-III schizophrenia (%) | 12 | 28 | 18.68* |
| DSM-III personality disorder (%) | 55 | 28 | 41.12* |
| Procriminal values (%) | 50 | 29 | 13.61* |
| Attitude unfavorable to convention (%) | 48 | 29 | 21.03* |

NOTE: Numbers under t or χ^2 are significant ($p < .05$) ts, $df > 100$, for continuous variables, or χ^2 (1, $N > 500$) for dichotomous variables for the comparison between entries in the first two columns. Asterisks indicate significant differences after a Bonferroni correction.

a. Numbers under these columns are means (accompanied by SDs) for continuous variables and percentages for dichotomous variables (indicated by %).

were used in the present study to ensure that the resulting instrument would be applicable not just to forensic psychiatric patients but to serious offenders in general who might be referred to forensic clinicians for appraisals of dangerousness and/or pretrial clinical assessments. Men who had no opportunity to recidivate were dropped from the study, leaving a final sample of 618.

The data presented in Table 1 show that the subjects were a group of serious offenders; all had been charged with a serious criminal offense and 85% had been charged with at least one violent offense (the current offense, a prior offense, or both). Of the remaining 15%, 6% had well-documented acts of interpersonal violence without criminal charges. In addition, 3% had been charged with arson, 2% with robbery, and 3% with weapons offenses (all of which would qualify as violence in some studies). Absent from this sample were offenders whose only crimes involved drug offenses, drunken driving, and property and economic offenses. Table 1 shows that many subjects had failed on prior conditional release or had escaped from custody; many had committed homicide; most had used a weapon in their index (i.e., current) offense; many were assessed as having antisocial, procriminal values; most qualified for a psychiatric diagnosis of person-

ality disorder or, less commonly, schizophrenia; and many had a serious alcohol abuse problem. Consequently, any statistical prediction instrument resulting from analysis of this sample cannot be expected to generalize to offenders with less serious criminal histories or, of course, to persons without any criminal conduct. However, because of the ways in which patients were referred to our institution (some for treatment, some for only a brief assessment), we believe that the sample was representative of offenders incarcerated for serious crimes.

Finally, it is important to note that the study variables (excluding recidivism) reflected only data available at the time the subject was referred to our institution. Some subjects remained for several years afterward to participate in treatment, whereas virtually all of the others were imprisoned for several years. The mean length of time between index offense and the point at which subjects first had an opportunity to recidivate was 5.31 years ($SD = 4.32$). The clinical and administrative characteristics of the study institution have been described elsewhere (Quinsey, 1981; Rice & Harris, 1993). Published data have indicated that treatment in the institution was at best weakly related to recidivism (Rice et al., 1992; Rice, Quinsey, & Houghton, 1990).

STUDY VARIABLES

All study variables (including the revised Psychopathy Checklist and *DSM-III* diagnosis), except those pertaining to recidivism, were coded retrospectively and exclusively from descriptive material contained in institutional files by a team of three research assistants who had extensive experience and training (mean greater than 3 years) as data coders for similar research projects. The clinical files have formed the basis for several other studies (e.g., Quinsey & Maguire, 1986; Rice & Harris, 1991). The institutional files included information from a variety of sources (e.g., psychosocial histories, information from other institutions, police reports, psychological test reports, questionnaires from patients' families). Most of the variables used in this study (shown in Table 1) are self-explanatory. However, a few require some additional explanation.

Elementary school maladjustment was rated on a 4-point scale from *none* (0) to *serious discipline and/or attendance problems* (3). Socio-

economic status was the highest rank order of either parent's occupation while the subject lived at home, according to the Blishen Scale (Blishen & McRoberts, 1976). Childhood behavior problems was the sum of items endorsed for the 12 problem behaviors noted before age 15 for a *DSM-III* diagnosis of antisocial personality disorder. Teen alcohol abuse was rated on a 4-point scale from *never drank* (0) to *serious drinking problem* (3). Separation from parents was scored as present if it occurred due to divorce, abandonment, or institutionalization before the subject turned 16. Adult aggression was rated on a 7-point scale from *no aggression* (1) to *occasional or frequent extreme aggression* (7). Level of Supervision Inventory was a modification of Andrews's (1982) 55-item scale (approximately 20% of the items were altered slightly). Property and violent offense history were summaries of criminal charges history using a modification (available from the authors) of Akman and Normandeau's (1967) seriousness scale. Victim injury was rated on a 7-point scale from *no injury* (1) to *death with mutilation* (7), following Quinsey and Chaplin (1982).

The revised 20-item version of the Psychopathy Checklist (PCL-R; Hare, 1991) was used to measure psychopathy. When individual items could not be coded (< 5%), scores were prorated. Although the PCL-R is often completed with an interview, scores in this study were based entirely on file information. Elsewhere, we have presented data showing that this method yields acceptable predictive validity (Harris et al., 1991) and the same psychometric properties (Harris, Rice, & Quinsey, in press) as the conventional method. The coding of diagnosis, including personality disorder, employed *DSM-III* criteria applied to file data available at the time of admission; it was not based on the diagnosis made by hospital physicians. Attitude supportive of crime was measured using an item from the Level of Supervision Inventory indicating procriminal, antisocial values (Andrews, 1982).

A separate team of three research assistants gathered all the outcome information and sent it for coding to the first team. Outcome data were obtained from the files of the coroner's office, the Lieutenant Governor's Board of Review (which maintains information about every insanity acquittee in Ontario), the Royal Canadian Mounted Police (a national database of criminal arrests and convictions including INTERPOL reports), the National Parole Service of Canada, and

provincial correctional and parole systems. Childhood history, adult adjustment, offense variables, and assessment variables were coded by the first team before recidivism data were obtained and coded to prevent inadvertent contamination of the historical variables by raters' knowledge of recidivism.

In coding recidivism, subjects were classified as violent failures if they incurred any new charge for a criminal offense against persons or were returned to a maximum security institution for violent behavior against persons that, in the judgment of the first team of raters, would otherwise have resulted in a criminal charge for an offense against persons. Violent offenses included all assaults and sexual assaults, armed robbery, forcible confinement, threatening, and pointing a firearm, but did not include possession of a weapon, robbery, or arson. We employed a dichotomous outcome variable (violent recidivism or not) because simple dichotomies of outcome have been found to perform as well as more sophisticated methods among correctional populations (Wormith & Goldstone, 1984). A subject was deemed to have an opportunity to recidivate when he was released to the street or was placed in a halfway house or an open psychiatric ward.

INTERRATER RELIABILITY

Twenty subjects were randomly chosen for an interrater reliability check. One rater (the one who coded the largest number of subjects) rated all 20 of the subjects selected for the reliability check, and each of the other two raters rated 10. Mean Pearson correlation coefficients were computed for continuous variables, and kappa (a statistic not subject to bias due to low base rates) for categorical variables (including recidivism). For all variables, the reliability criterion was set at .70, and variables not reaching this criterion were dropped from the study. For all variables retained, the mean correlation coefficient was .90 and the mean kappa statistic was .83.

ACTUARIAL PREDICTION

We sought to derive a prediction instrument that would work well for serious offenders in general and that used a variety of predictors applicable to a variety of offenders. Approximately one third of the

present subjects were men who had been acquitted on a criminal offense because of insanity. The data were subjected to preliminary analyses to determine whether the relationships between study variables and violent recidivism were different for the insanity acquittees and the other men, and they were not. Thus, in the remainder of the present report, data from the insanity acquittees were combined with those of the other subjects to derive a statistical prediction instrument. A multivariate approach was employed to select variables for inclusion in the prediction instrument. To select a reduced set of variables to be used in the final analysis, a separate stepwise discriminant analysis¹ (α to enter and remove = .25) using each of the four sets of variables shown in Table 1 (childhood history, adult adjustment, index offense, and assessment results) was conducted for each of five groups of subjects who had an opportunity to recidivate. The five groups were: (a) the entire sample ($N = 618$), (b) the treated subjects ($n = 332$), (c) the assessed subjects ($n = 286$), (d) a randomly selected half of the entire sample, and (e) the remaining half of the entire sample. Based on order of entry in stepwise analyses, up to four variables were selected from each variable set for each group. Because stepwise multivariate results are often unstable, only those variables (of the 19 candidate variables) that were selected by at least three of the five analyses of the different subject groups were allowed to enter into the final stepwise multiple discriminant analysis.

RESULTS

At the last follow-up, in April 1988, the average time at risk was 81.5 months ($SD = 60.6$). A total of 191 (31%) of the at-risk subjects were violent failures. Table 2 shows the univariate comparisons between those subjects who exhibited violent recidivism and those who did not. Because of the number of comparisons, the Bonferroni correction procedure was applied ($\alpha = .001$).

The final discriminant analysis using the entire sample identified 12 variables for inclusion in the final statistical prediction instrument. These variables and their relationship with violent recidivism for the entire sample are indicated in Table 2. The multiple correlations for a final discriminant function using the 12 variables selected in the

TABLE 2: Univariate Correlations Between Each Final Predictor Variable and Violent Recidivism Plus Multiple Correlations Achieved With the Addition of Each

| <i>Variable</i> | <i>r^a</i> | <i>R^b</i> |
|--------------------------------------|----------------------|----------------------|
| Psychopathy Checklist | .34 | — |
| Separation from parents under age 16 | .25 | .406 |
| Victim injury in index offense | -.16 | .429 |
| <i>DSM-III</i> schizophrenia | -.17 | .439 |
| Never married | .18 | .446 |
| Elementary school maladjustment | .31 | .450 |
| Female victim-index offense | -.11 | .454 |
| Failure on prior conditional release | .24 | .456 |
| Property offense history | .20 | .457 |
| Age at index offense | -.26 | .458 |
| Alcohol abuse history | .13 | .459 |
| <i>DSM-III</i> personality disorder | .26 | .459 |

a. For all *r*s, $p < .05$.

b. For all *R*s, $p < .0001$.

analysis using the entire sample but applied to each of the subsamples are shown in Table 3.

Although the multiple discriminant approach as used above is sound, multivariate functions generally perform much less well on cross-validation (i.e., when used on a new sample from the same population). Part of the reason for this has to do with the fact that when there are many potential variables considered for inclusion, some that enter into the final equation happen by chance to be important for the particular sample selected, but would not turn out to be important in a new sample. We partially accounted for this possibility by eliminating from consideration those variables that added to the multiple *R* for two or fewer of the subsamples considered in this study. However, another reason that multiple discriminant solutions tend to shrink on cross-validation has to do with the weighting of the variables. In the final discriminant function, the weights of each of the selected variables is maximized so as to provide the best possible fit for the particular sample used. Just as in the selection of variables, there is a certain amount of capitalization on chance in the selection of the best weights. Therefore, to reduce the amount of shrinkage that would be expected on cross-validation due to the calculation of weights, we used two different approaches.

In the first procedure, we assigned unitary weights to each of the selected variables. That is, we constructed an equation that forced each predictor to participate equally. This approach yields a liberal estimate of the effects of predictive shrinkage that would occur on cross-validation (Tatsuoka, 1988). We did this by first standardizing each variable (i.e., converting each to a z score with a mean of 0 and a standard deviation of 1). Then, for each subject, a score was obtained by summing his standard scores for each variable. The sum of these variables was significantly correlated, $r(618) = .44$, with violent recidivism. The performance of this unitary weight equation (in terms of the correlation between total score and violent recidivism) with each of the construction samples is shown in Table 3. The entire range of scores on the unitary weight equation was divided into nine steps or "bins" of equal size; the rate of violent recidivism for subjects at each step is shown in the top panel of Figure 1. Also shown are the number of subjects assigned to each of the nine steps. It may be seen that the equation performed well, in that it assigned subjects to all levels of risk and that the probability of violent recidivism increased steadily from near 0 to 100% as scores on the prediction equation increased.

The other method of using the 12 variables to construct a prediction equation was described by Nuffield (1982). According to this method, predictor variables are selected based on univariate correlations with violent recidivism and then are weighted in proportion to their abilities to discriminate recidivists from nonrecidivists. Although the 12 variables we used were selected by a multivariate rather than univariate procedure, all were significant at least at the .05 level before a Bonferroni correction was applied. The value of using the variables selected by the multivariate procedure was that it reduced the number of variables for inclusion in the prediction in a way that eliminated redundant variables. Then, following Nuffield (1982), the recidivism rates for each value (or range of values) of each of the 12 predictor variables were determined. For every difference (plus or minus) of more than 5% from the mean overall violent recidivism rate (31%), a weighting of one was added. Thus, for example, for the variable "never married," subjects who had married had a violent recidivism rate of 21% (two full increments of 5% below the mean), and those who had not married had a recidivism rate of 38% (one full increment of 5% above the mean). Thus, for that variable, a subject who had married

TABLE 3: Correlations Between the Final 12 Predictor Variables and Violent Recidivism for Multiple Discriminant Function (weighted) and the Final Statistical Instrument Using Unitary Weights (unweighted)

| <i>Sample</i> | <i>Weighted</i> | | <i>Unweighted</i> |
|------------------------|-----------------|----------------------|----------------------|
| | <i>n</i> | <i>R^a</i> | <i>r^a</i> |
| All subjects | 618 | .459 | .424 |
| Treated subjects | 332 | .527 | .462 |
| Subjects assessed only | 286 | .454 | .370 |
| 1st random subjects | 307 | .522 | .477 |
| 2nd random subjects | 311 | .432 | .370 |

a. For all correlations, $p < .001$.

would obtain a score of -2 , whereas a subject who had never married would obtain a score of $+1$. The variable with the highest possible range of scores using this method was the Psychopathy Checklist (because it had the highest correlation with violent recidivism), and possible scores on this variable ranged from -5 to $+12$. Using this method with each of the 12 final variables resulted in a risk score that ranged between -27 and $+35$. (A copy of the final scoring key can be obtained from the authors.) The correlation between scores on this prediction equation and violent recidivism was .44, identical to the correlation obtained using the unitary weight method described above. The violent recidivism rates for nine equal-size bins, each encompassing a range of 8 points, are illustrated in the middle panel of Figure 1. Again, it may be seen that the Nuffield equation performed well in terms of assigning subjects to risk levels, and that actual violent recidivism increased steadily with predicted risk level. Indeed, because the performance of this equation was at least as good as the unitary weight version, and because this method is easier to apply to new samples, we decided to adopt this equation based on the Nuffield method as our prediction instrument.

Because the set of 12 predictor variables contains information that is often unavailable without a detailed history and specialized diagnostic procedures, it is of interest to know how well violent recidivism could have been predicted from routinely available information. When the PCL-R, *DSM-III* diagnosis, and childhood variables (all of which require detailed psychosocial histories) were dropped, the Nuffield procedure applied to the remaining seven variables yielded a score

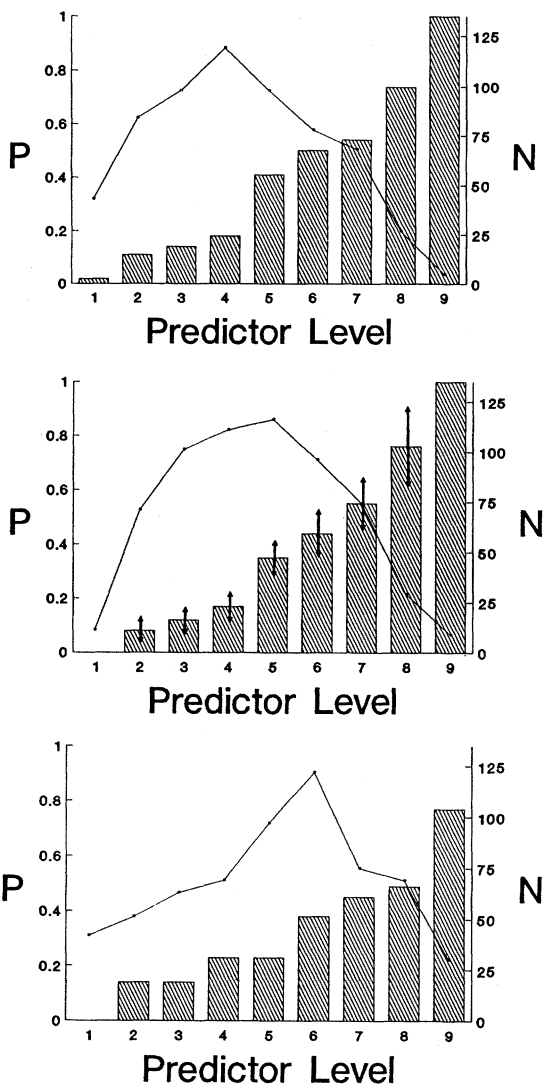


Figure 1: The performance of the unitary weight statistical prediction equation using z scores (top panel), the final statistical instrument using the Nuffield method to weight variables (middle panel), and the final instrument using a reduced set of seven variables (bottom panel).

NOTE: Bars indicate the probability of violent recidivism (P) for subjects at each of nine equal size steps. The solid line indicates the number of subjects (N) in each step or bin. Arrows in the middle panel indicate the 95% confidence intervals for the violent recidivism rates for subjects in each bin.

that had a correlation of $r(618) = .36$ with violent recidivism. The performance of this more limited equation is summarized in the bottom of Figure 1. It can be seen that, although still adequate, there was no subgroup in which all members reoffended violently.

PSYCHOMETRIC PROPERTIES

The mean score on the final actuarial instrument was .91 ($SD = 12.9$). If a prediction instrument is to be used in making decisions about individual cases, issues of measurement error should be addressed. The first such issue pertains to the standard error of measurement (SEM; i.e., the relationship between a subject's hypothetical "true" score and his actual obtained score). The SEM is a function of both the standard deviation and the reliability of a measure. The reliability of the final prediction instrument using the Nuffield method was evaluated by comparing final scores based on the two independent coders described above. The resulting interrater reliability was a Pearson correlation coefficient of .90. Thus the SEM for the instrument was equal to $SD\sqrt{1-r} = 4.1$. This SEM is small enough to ensure that an individual subject's true score lay close to his obtained score.

The second issue pertains to the confidence one may place in the obtained violent failure rates for the samples of subjects in each bin shown in the middle panel of Figure 1. The magnitudes of such confidence intervals depend on the degree of confidence desired, the variability of violent recidivism, and the number of subjects in the bin. The 95% confidence intervals for each bin are shown in the figure. Most noteworthy is the fact that most adjacent confidence intervals do not show great overlap. Thus the data presented here suggest that the final prediction instrument possesses acceptable psychometric properties. That is, the levels of relative risk identified by the instrument are discriminably different from each other. Also, the standard error of measurement possessed by the instrument suggests that individual subjects can be assigned reliably to levels of risk.

Of course, a crucial psychometric issue is predictive validity. Table 4 presents the accuracy with which violent recidivism was predicted (using the final instrument) in the entire sample for each of several cutoff scores. The results illustrate the usual trade-off between sensitivity (i.e., the proportion of violent recidivists identified) and speci-

TABLE 4: Predictive Accuracy of the Final Instrument at Varying Criteria for a Prediction of Violent Recidivism

| <i>Cutoff Score</i> | <i>Percentage Correct</i> | <i>False Positives</i> | <i>Sensitivity</i> | <i>Specificity</i> | ϕ^a |
|---------------------|---------------------------|------------------------|--------------------|--------------------|----------|
| 20th percentile | 47 | .64 | .95 | .25 | .24 |
| 40th percentile | 61 | .57 | .87 | .49 | .34 |
| 60th percentile | 70 | .49 | .72 | .69 | .38 |
| Base rate | 73 | .45 | .60 | .78 | .37 |
| 80th percentile | 74 | .39 | .41 | .88 | .33 |

a. For all phi coefficients, $p < .001$.

ficity (i.e., the proportion of nonrecidivists correctly identified) obtained with any test. They indicate that, assuming approximately an equivalent cost of misses (i.e., false negatives) and false alarms (i.e., false positives), good performance was obtained when the cutoff score was set so that the selection rate was equal to the base rate of violent recidivism. This cutoff corresponded to the boundary between the fifth and sixth bins in the middle panel of Figure 1. At that cutoff score, total accuracy was almost maximized whereas both a low false positive rate and reasonable sensitivity were maintained. Examination of the binomial effect size (the likelihood of violence both above and below the cutoff; Rosenthal, 1990) for cutoff scores around the base rate reveals that subjects above the cutoff exhibited more than twice the rate of violent recidivism (69%) than subjects below the cutoff (31%).

DISCUSSION

The multiple discriminant analyses using 12 predictor variables yielded a multiple R of .459 in the whole sample and multiple R s ranging from .432 to .527 in the various subsamples. When each variable was allowed to participate equally in the prediction (done by transforming scores on all variables to z scores), the correlations with violent recidivism ranged from .370 to .477. The application of the Nuffield (1982) method (which is much easier to use with new samples) resulted in an equation with performance identical to that obtained using the z -score method. These values translate into a

classification accuracy of about 75%. These modest decrements in performance obtained using the *z*-score or Nuffield methods rather than using the weightings obtained from the multivariate analyses are very close to what might be expected if full cross-validations using the multivariate equation were conducted with independent but similar samples of offenders (Kerlinger & Pedhazur, 1973). A reduced equation using only seven easily obtained variables showed some decrement in performance but still exhibited some ability to discriminate between subjects who recidivated and those who did not.

It should be noted that recidivism could be evaluated only for those offenders who had been released. Preliminary comparisons between the 618 subjects and 67 other men who qualified as subjects but were never released revealed a few significant differences. The largest difference between the released and retained groups was in the seriousness of the index offense. Although this variable was negatively related to recidivism, the retained group had higher seriousness of index offense scores. Similarly, degree of victim injury and having a female victim were negatively related to violent recidivism, but retained offenders were more likely to have caused serious injury and to have had a female victim than released subjects. The only significant difference between released and retained subjects that, according to our prediction instrument, indicated that the retained subjects would have had higher rates of violent recidivism was that fewer of the retained subjects had been married. Because there were so few unreleased subjects, we believe that the prediction instrument would have performed equally well if all 685 potential subjects had had an opportunity to reoffend.

The issue of generalization to other populations warrants further comment. At first glance it might seem that, because all had been psychiatric patients, some were psychotic, and some were found insane, the present subjects were very different from criminal offenders in general. Without doubt, the present subjects comprised only dangerous offenders, each of whom had already committed at least one serious offense. It is not at all clear, however, that the present subjects differed materially from serious offenders found in prison. There are several bases for this statement: First, the majority (55%) of the present sample (all of the assessed and many of the treated subjects) were not insane but had been convicted. Second, most (58%)

of the insanity acquittees were not psychotic but had been found insane based on psychiatric testimony that they had severe antisocial personality disorders. Third, although a significant minority (about 30%) of the present subjects did qualify for a *DSM-III* diagnosis of psychosis, recent studies (Hodgins, 1990) indicate that a surprisingly large proportion (as high as 25%) of prison inmates also qualify for *DSM-III* psychotic diagnoses. That proportion might be even higher among serious offenders. Fourth, questions about the comparability of serious offenders among prison inmates and forensic psychiatric patients are very difficult to evaluate because very few investigators have published data on the demographic, social, criminal, or psychiatric histories of serious offenders from either population. We have reported elsewhere (Harris, Rice, Quinsey, Chaplin, & Earls, 1992) that sex offenders in our population and sex offenders in a Canadian prison were indistinguishable on variables reflecting criminal, demographic, and social history. Fifth, most of the present findings pertaining to individual predictor variables (age, marital status, psychopathy, criminal history, failure on prior conditional release, alcohol use, antisocial conduct in childhood) and recidivism have been reported by other investigators studying offender populations exclusively in prisons. In summary, we believe there are good grounds to anticipate that the present results will generalize both to mentally disordered offenders from other jurisdictions *and* to serious offenders in prison populations. Of course, further research is required to test our expectation.

The prediction equation used only static linear predictive relationships and did not include interactions. Results from our earlier work (Harris et al., 1991), for example, suggested that age and psychopathy interacted in their effects on violent recidivism such that, among nonpsychopaths (i.e., those scoring less than 25 on the PCL-R), rates of violent recidivism declined with age whereas among psychopaths, there was no such decline. Further research is required to incorporate such interactive variables into the statistical prediction of violent recidivism.

Another task of future research is to establish prediction equations for more homogeneous subgroups of offenders that have high enough base rates of violent offending to permit a meaningful increase in predictive accuracy. For example, several investigators have reported that child molesters with male victims are at greater risk to recidivate than those with only female victims (Furby, Weinrott, & Blackshaw,

1989; Quinsey, 1986; Rice, Quinsey, & Harris, 1991) and that sex offenders with more deviant sexual preferences are more likely to fail (Rice, Harris, & Quinsey, 1990; Rice et al., 1991). Thus a prediction equation for sex offenders might profitably employ additional aspects of offense history and phallometric data as predictors. Similarly, different variables may be necessary to accurately predict violent recidivism among arsonists compared to violent robbers compared to spouse murderers.

The level of accuracy attained by the actuarial model raises the issue of its practical applications, of which there are several. Actuarial models can be used in such policy decisions as the determination of the level of perimeter security required to safely contain offenders within an institution. With respect to treatment, actuarial predictors of risk can be used to identify groups of offenders for whom specialized treatment or supervision are required. For example, it has been suggested that treatments designed to alter risk of recidivism are more effective when directed toward higher- rather than lower-risk offenders (Andrews, 1989).

How might statistical predictions be used in making release and disposition decisions about individual offenders? Because the actuarial instrument developed in the present study was not perfectly accurate, and recognizing that each individual is unique, one might be tempted to ignore statistical predictions and base decisions solely on clinical judgment and intuition. However, such a strategy would inevitably result in more errors (greater risk to the public, more unnecessary incarceration, or both) than would the use of an actuarial instrument. That is, the data on the prediction of violence indicate that clinicians are insensitive to variations in the base rate of violent offending, show poor agreement among themselves, make the same judgments as laypersons, and are less accurate than actuarial models (Faust & Ziskin, 1988; Quinsey & Ambtman, 1979; Quinsey & Maguire, 1986), a series of findings that is consistent with a much larger literature on human judgment in probabilistic situations.

Clinical judgment can be improved, however, through the use of actuarial information; this has been referred to as "structuring discretion" (Gottfredson, Wilkins, & Hoffman, 1978). In this approach to decision making about an individual, an actuarial estimate of risk is used to anchor clinical judgment. More specifically, clinicians can use

dynamic (changeable) information such as progress in treatment, change in procriminal attitudes, and the amount and quality of supervision in the postrelease environment to adjust the risk level computed by the actuarial prediction instrument. If adjustments are made conservatively and *only* when a clinician believes, on good evidence, that a factor is related to the likelihood of violent recidivism in an individual case, predictive accuracy may be improved.

What about offenders who are classified as high risk but who would not in fact recidivate upon release (i.e., the false positives)? Because being a high risk does not mean that a person will necessarily commit a new violent offense, decisions should be made that protect the public while reducing the human cost of false positives (Quinsey & Walker, 1992). We believe that there are a few offenders whose offenses are so serious and their actuarial risk scores so high that they are proper candidates for incapacitation (through indefinite incarceration). Of course, it is not for researchers or clinicians, but rather for the courts and other judicial bodies, to decide how serious the history and how high the risk must be before incapacitation is imposed. Otherwise, actuarial measures of risk are best suited for ensuring that more intensive treatments and more rigorous supervisory methods are employed with the offenders who need them the most, those of high risk.

NOTE

1. It might be argued that logistic regression would have been the more appropriate multivariate technique. We performed all of the analyses with logistic regression, and the variables selected and their order were always identical to the results from the multiple discriminant analyses. We preferred multiple discriminant analyses because they readily yield *Rs* and permit evaluation of the proportions of shared variance and because we did not intend to use the beta weights from either analysis.

REFERENCES

- Akman, D. D., & Normandeau, A. (1967). The measurement of crime and delinquency in Canada: A replication study. *British Journal of Criminology*, 7, 129-149.
- Andrews, D. A. (1982). *The Level of Supervision Inventory*. Toronto: Ministry of Correctional Services (Ontario).

- Andrews, D. A. (1989). Recidivism is predictable and can be influenced: Using risk assessments to reduce recidivism. *Forum on Corrections Research*, 1, 11-18.
- Blishen, B. R., & McRoberts, H. A. (1976). A revised socioeconomic index for occupations in Canada. *Canadian Review of Sociology and Anthropology*, 13, 71-79.
- Faust, D., & Ziskin, J. (1988). The expert witness in psychology and psychiatry. *Science*, 241, 501-511.
- Furby, L., Weinrott, M. R., & Blackshaw, L. (1989). Sex offender recidivism: A review. *Psychological Bulletin*, 105, 3-30.
- Gottfredson, S. D. (1987). Statistical and actuarial considerations. In F. Dutilleul & C. Foust (Eds.), *The prediction of violence* (pp. 71-81). Springfield, IL: Charles C Thomas.
- Gottfredson, D. M., Wilkins, L. T., & Hoffman, P. B. (1978). *Guidelines for parole and sentencing: A policy control method*. Toronto: Lexington Books.
- Hare, R. D. (1991). *The Revised Psychopathy Checklist*. Toronto: Multi-Health Systems.
- Harris, G. T., Rice, M. E., & Cormier, C. (1991). Psychopathy and violent recidivism. *Law and Human Behavior*, 15, 625-637.
- Harris, G. T., Rice, M. E., & Quinsey, V. L. (in press). Psychopathy as a taxon: Evidence that psychopaths are a discrete class. *Journal of Consulting and Clinical Psychology*.
- Harris, G. T., Rice, M. E., Quinsey, V. L., Chaplin, T. C., & Earls, C. (1992). Maximizing the discriminant validity of phallometric assessment. *Psychological Assessment*, 4, 502-511.
- Hodgins, S. (1990). Prevalence of mental disorders among penitentiary inmates in Quebec. *Canada's Mental Health*, 38, 1-4.
- Holland, T. R., Holt, N., Levi, M., & Beckett, G. E. (1983). Comparison and combination of clinical and statistical predictions of recidivism among adult offenders. *Journal of Applied Psychology*, 68, 203-211.
- Kerlinger, F. N., & Pedhazur, E. J. (1973). *Multiple regression in behavioral research*. New York: Holt, Rinehart & Winston.
- Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques*. Beverly Hills, CA: Sage.
- Monahan, J., & Steadman, H. (Eds.). (in press). *Violence and mental disorder: Developments in risk assessment*. Chicago: University of Chicago Press.
- Nuffield, J. (1982). *Parole decision-making in Canada: Research towards decision guidelines*. Ottawa: Ministry of Supply and Services Canada.
- Quinsey, V. L. (1980). The base rate problem and the prediction of dangerousness: A reappraisal. *Journal of Psychiatry and Law*, 8, 329-340.
- Quinsey, V. L. (1981). The long term management of the mentally disordered offender. In S. J. Hucker, C. D. Webster, & M. Ben-Aron (Eds.), *Mental disorder and criminal responsibility* (pp. 137-155). Toronto: Butterworths.
- Quinsey, V. L. (1984). Politique institutionnelle de liberation: Identification des individus dangereux: une revue de la literature (Institutional release policy and the identification of dangerous men: A review of the literature). *Criminologie*, 17, 53-78.
- Quinsey, V. L. (1986). Men who have sex with children. In D. Weisstub (Ed.), *Law and mental health: International perspectives* (pp. 140-172). New York: Pergamon.
- Quinsey, V. L., & Ambtman, R. (1979). Variables affecting psychiatrists' and teachers' assessments of the dangerousness of mentally ill offenders. *Journal of Consulting and Clinical Psychology*, 47, 353-362.
- Quinsey, V. L., & Chaplin, T. C. (1982). Penile responses to nonsexual violence among rapists. *Criminal Justice and Behavior*, 9, 312-324.
- Quinsey, V. L., & Maguire, A. (1986). Maximum security psychiatric patients: Actuarial and clinical prediction of dangerousness. *Journal of Interpersonal Violence*, 1, 143-171.

- Quinsey, V. L., & Walker, W. D. (1992). Dealing with dangerousness: Community risk management strategies with violent offenders. In R. DeV. Peters, R. J. McMahon, & V. L. Quinsey (Eds.), *Aggression and violence throughout the life span* (pp. 244-262). Newbury Park, CA: Sage.
- Rice, M. E., & Harris, G. T. (1991). Firesetters admitted to a maximum security psychiatric institution: Characteristics of offenders and offenses. *Journal of Interpersonal Violence*, 6, 461-475.
- Rice, M. E., & Harris, G. T. (1992). A comparison of criminal recidivism among schizophrenic and nonschizophrenic offenders. *International Journal of Law and Psychiatry*, 15, 397-408.
- Rice, M. E., & Harris, G. T. (1993). Ontario's maximum security hospital at Penetanguishene: Past, present and future. *International Journal of Law and Psychiatry*, 16, 195-215.
- Rice, M. E., Harris, G. T., & Cormier, C. (1992). Evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior*, 16, 399-412.
- Rice, M. E., Harris, G. T., Lang, C., & Bell, V. (1990). Recidivism among male insanity acquittees. *Journal of Psychiatry and Law*, pp. 379-403.
- Rice, M. E., Harris, G. T., & Quinsey, V. L. (1990). A followup of rapists assessed in a maximum security psychiatric facility. *Journal of Interpersonal Violence*, 5, 435-448.
- Rice, M. E., Quinsey, V. L., & Harris, G. T. (1991). Sexual recidivism among child molesters released from a maximum security institution. *Journal of Consulting and Clinical Psychology*, 59, 381-386.
- Rice, M. E., Quinsey, V. L., & Houghton, R. (1990). Predicting treatment outcome and recidivism among patients in a maximum security token economy. *Behavioral Sciences and the Law*, 8, 313-326.
- Rosenthal, R. (1990). How are we doing in soft psychology? *American Psychologist*, 45, 775-777.
- Steadman, H. J. (1983). Predicting dangerousness among the mentally ill: Art, magic and science. *International Journal of Law and Psychiatry*, 6, 381-390.
- Steadman, H. J. (1987). How well can we predict violence for adults? A review of the literature and some commentary. In F. Dutile & C. Foust (Eds.), *The prediction of criminal violence* (pp. 5-19). Springfield, IL: Charles C Thomas.
- Tatsuoka, M. M. (1988). *Multivariate analysis*. New York: Macmillan.
- Wormith, J. S., & Goldstone, C. S. (1984). The clinical and statistical prediction of recidivism. *Criminal Justice and Behavior*, 11, 3-34.